ELSEVIER

Historical review: Deciphering the genetic code – a personal account

Marshall Nirenberg

Laboratory of Biochemical Genetics, National Heart, Lung and Blood Institute, National Institutes of Health, 9000 Rockville Pike, MSC – 1654, Building 10, Room 7N-315, Bethesda, MD 20892-1654, USA

This is an autobiographical description of the events that led to the breaking of the genetic code and the subsequent race to decipher the code. The code was deciphered in two stages over a five-year period between 1961 and 1966. During the first stage, the base compositions of codons were deciphered by the directing cell-free protein synthesis with randomly ordered RNA preparations. During the second phase, the nucleotide sequences of RNA codons were deciphered by determining the species of aminoacyl-tRNA that bound to ribosomes in response to trinucleotides of known sequence. Views on general topics such as how to pick a research problem and competition versus collaboration also are discussed.

I would like to tell you how the genetic code was deciphered from a personal point of view. I came to the National Institutes of Health (NIH) in 1957 as a post-doctoral fellow with Dewitt Stetten, Jr, a wise, highly articulate scientist and administrator, immediately after obtaining a PhD in biochemistry from the University of Michigan in Ann Arbor. The next year, I started work with William Jakoby and, by enrichment culture, I isolated a *Pseudomonad* that grew on γ-butyrolactone and purified three enzymes involved in the catabolism of γ-hydroxybutyric acid [1].

There was a weekly seminar in Stetten's laboratory in which Gordon Tomkins (Figure 1), who worked in a different laboratory, participated. Gordon was brilliant, with a wonderful associative memory and a magnificent sense of humor. His seminars were superb, especially his description of the step-by-step developments in the problem that he intended to discuss. Towards the end of my post-doctoral fellowship, Gordon replaced Herman Kalckar as head of the Section of Metabolic Enzymes and offered me a position as an independent investigator in his laboratory. The other independent investigators in the laboratory were Elizabeth Maxwell and Victor Ginsberg, who were carbohydrate biochemists, and Todd Miles, a nucleic-acid biochemist. It was a wonderful opportunity and I decided then that if I was going to work this hard I might as well have the fun of exploring an important problem.

In my opinion, the most exciting work in molecular biology in 1959 were the genetic experiments of Monod and Jacob on the regulation of the gene that encodes β-galactosidase in *Escherichia coli* and that the mechanism of protein synthesis was one of the most exciting areas in biochemistry. Some of the best biochemists in the world were working on cell-free protein synthesis, and I had no experience with either gene regulation or protein synthesis, having previously worked on sugar transport, glycogen metabolism and enzyme purification. After thinking about this for a considerable time, I finally decided to switch fields. My immediate objective was to investigate the existence of mRNA by determining whether cell-free protein synthesis in *E. coli* extracts was stimulated by an RNA fraction or by DNA. In the longer term, my objective was to achieve the cell-free synthesis of penicillinase, a small inducible enzyme that

Figure 1. Gordon Tompkins. Gordon was brilliant, highly articulate and very funny. He was a charismatic individual who created a stimulating atmosphere and encouraged exploration. In 1958, towards the end of my post-doctoral fellowship at the NIH, he offered me a position as an independent investigator in his laboratory.

Corresponding author: Marshall Nirenberg (mnirenberg@nih.gov).

http://tibs.trends.com 0968-0004/$ - see front matter © 2003 Elsevier Ltd. All rights reserved. doi:10.1016/j.tibs.2003.11.009

lacks cysteine so that I could explore mechanisms of gene regulation. I thought that in the absence of cysteine the synthesis of penicillinase might proceed, whereas synthesis of most other proteins might be reduced.

In England, Pollock [2] had shown that penicillinase is inducible in *Bacillus cerus* and had isolated mutants that differed in the regulation of the penicillinase gene. In 1959, tRNA was recently discovered but mRNA was unknown. At that time, the only clues that RNA might function as a template for protein synthesis were a report by Hershey *et al.* [3], showing that a fraction of RNA is synthesized and degraded rapidly in *E. coli* infected with T2 bacteriophage, and a paper by Volkin and Astrachan [4], which showed that infection of *E. coli* by T2 bacteriophage resulted in the rapid turnover of a fraction of RNA that had the base composition of bacteriophage rather than the DNA of *E. coli*. If mRNA did exist, I thought that it might be contained in ribosomes because amino acids were known to be incorporated into protein on these organelles. I estimated it would take me two years to set up a cell-free system to determine whether RNA or DNA stimulated protein synthesis, which was a pretty accurate estimate.

I knew this was a risky problem to work on because starting out as an independent investigator you are supposed to hit the deck running and prove that you are an effective, productive investigator. One evening I saw Bruce Ames working in his laboratory. Because I thought he was one of the best young scientists at the NIH I described my research plan and asked for his evaluation. He just looked at me and said 'It is suicidal'. Although we both agreed that it was a dangerous project to work on, I thought suicidal was a little extreme. On the one hand I wanted to explore an important problem, on the other I was afraid of failure, but the wish to explore was much greater than the fear of failure.

As soon as I moved to Gordon's laboratory I started to make cell-free extracts that incorporated amino acids into protein, and to prepare DNA and RNA from ribosomes of penicillinase inducible and constitutive strains of *B. cerus*. I devised a sensitive assay for penicillinase and starting with conditions that had been devised by Lamborg and Zamecnik and his colleagues [5], I tried to obtain the *de novo* synthesis of penicillinase following addition of either RNA or DNA fractions from either *B. cerus* or *E. coli*. Systematically, I explored the optimum conditions for cell-free synthesis and showed that RNA prepared from ribosomes of *B. cerus* that expressed penicillinase constitutively stimulated penicillinase synthesis by 10–15%, but RNA from either uninduced ribosomes or DNA had no effect. However, the stimulation of penicillinase synthesis was small and it was clear that I needed a more sensitive assay.

Usually around noon, Gordon Tomkins would come into my laboratory with a sandwich and we would go into the hall and talk about my work, his work, and various exciting results that had been published. I always stopped to talk to him, even though the extract that I was preparing was slowly dying in an ice bucket, because these were wonderful conversations. Gordon encouraged me and created an exciting atmosphere for young investigators.

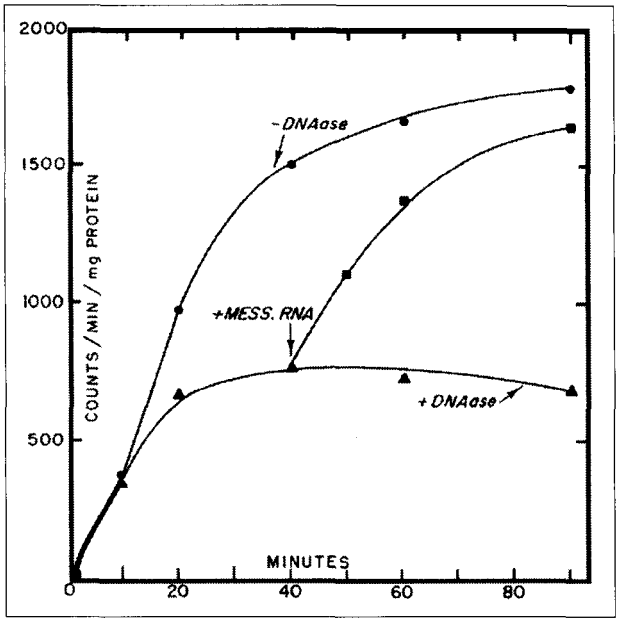

Figure 2. Incorporation of ^{14}C–labeled valine into protein in *Escherichia coli* extracts. Endogenous incorporation of radioactive amino acids into protein in *E. coli* extracts was high. However, amino acid incorporation ceased after incubation for ~40 min with DNase I. I found that I could freeze *E. coli* extracts and thaw them without loss of activity, so I incubated *E. coli* extracts in the absence of radioactive amino acids for 40 min, divided the extracts into small aliquots and froze them for use later in different experiments. Endogenous incorporation of radioactive amino acids was greatly reduced in such extracts, and addition of mRNA preparations from ribosomes clearly stimulated amino acid incorporation into protein [16,34,49]. Reproduced from Ref. [16].

After working on this for about for about a year and a half, Heinrich Matthei came to my laboratory as a post-doctoral fellow. Heinrich was a plant physiologist from Germany who was a post-doctoral fellow at Cornell who wanted to work on protein synthesis. He came under the impression that, because the NIH is such a big institution, many people would be working on protein synthesis. He stopped in Roy Vagelos's laboratory and Roy sent him to me because I was the only person at the NIH who was studying cell-free protein synthesis. We needed a more sensitive assay, so I suggested that Heinrich use the cell-free amino-acid-incorporating system that I had optimized to measure the incorporation of radioactive amino acids into protein. Heinrich insisted on preparing 20 ^{14}C-labeled amino acids by growing algae in the presence of ^{14}C-bicarbonate, hydrolyzing the protein and purifying each of the ^{14}C-labeled amino acids, because this is what he had done previously.

Using this more sensitive assay it was immediately apparent that RNA from ribosomes, but not DNA, stimulated incorporation of radioactive amino acids into protein [6,7]. I jumped for joy because this was the first definitive demonstration *in vitro* that mRNA existed and was required for protein synthesis. We fractionated RNA from ribosomes and found, as expected, that only a small portion stimulated amino acid incorporation into protein [8].

We made three trivial technical advances that had a tremendous effect on our work. First, I established conditions that enabled us to freeze and thaw *E. coli* extracts with little or no loss in the ability to incorporate

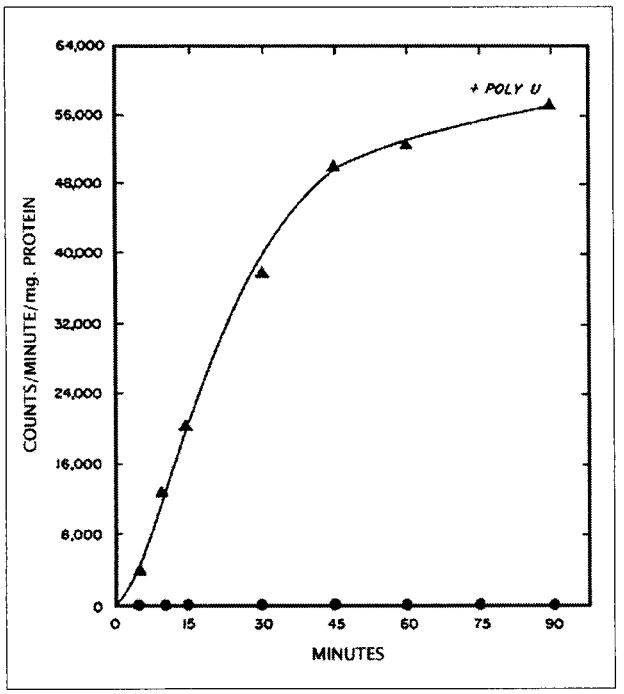

Figure 3. Poly(U) greatly stimulates the incorporation of radioactive phenylalanine into poly-phenylalanine [8].

amino acids into protein. Second, as shown in Figure 2, basal, endogenous incorporation of ^{14}C-labeled valine into proteins in the absence of mRNA was high, hence, the increase in amino acid incorporation caused by mRNA was relatively small. We confirmed the reports of Kameyama and Novelli [9] and Tissieres *et al.* [10] that DNase I inhibited the incorporation of amino acids into protein in cell-free *E. coli* extracts. Therefore, we incubated *E. coli* extracts in the presence of DNase I but without a radioactive amino acid for 40 min until endogenous amino acid incorporation had almost stopped [16,34,49]. Then we divided the extracts were into small portions and froze them for use later. Because the endogenous incorporation of radioactive amino acids into protein was low in these extracts, it was stimulated markedly by the addition of mRNA. Third, the standard method of washing radioactive protein precipitates in trichloracetic acid to remove radioactive amino acids involved repeated centrifugation and resuspension of protein pellets, which was very laborious and time consuming. One evening I compared this standard method with washing protein precipitates by filtration through Millipore filters. The results were identical. By using frozen–thawed, preincubated *E. coli* extracts and washing radioactive precipitates on Millipore filters we could do as much as in one day as had previously taken us 8–10 days.

I then obtained yeast rRNA and tobacco mosaic virus (TMV) RNA and we found that both were as active as mRNA. However, RNA from TMV was 30–50 times more active than ribosomal RNA at stimulating amino acid incorporation into protein. I called Heinz Frankel-Conrat in Berkeley, a world expert on TMV who had a mutant with an amino acid replacement in the viral coat protein, to tell him our results. He invited me to come to his laboratory to synthesize radioactive protein directed by RNA from wild-type and mutant TMV, with the intention that he and a colleague would purify and characterize the products to determine whether the radioactive protein synthesized was TMV coat protein. I felt like Marco Polo exploring a new area.

Before going to Frankel-Conrat's laboratory I obtained some poly(U) and instructed Heinrich to make 20 different solutions, each with 19 cold amino acids and one radioactive amino acid, to detect poly(U)-dependent incorporation of a single radioactive amino acid into protein. After working in Frankel-Conrat's laboratory for about a month, Heinrich called me very excitedly to tell me that poly(U) was extraordinarily active in stimulating the incorporation of only phenylalanine into protein [8] (Figure 3). I immediately returned to Bethesda. We also showed that single-stranded poly(U) functions as mRNA, but double-stranded or triple-stranded poly(U)–poly(A) helices do not [8] (Figure 4). This was the first RNA antisense experiment. In addition, we showed that poly(C) directs the incorporation only of proline into protein [8].

I thought the poly(U) result wouldn't be believed unless we characterized the radioactive polyphenylalanine product of the reaction very carefully. As shown in Table 1, hydrolysis of the ^{14}C-labeled polyphenylalanine by HCl recovered stoichiometric amounts of ^{14}C-labeled phenylalanine. I also thought we should show that the solubility of the ^{14}C-polyphenylalanine was the same as that of authentic polyphenylalanine, but because I knew nothing about this I went to Chris Anfinson's laboratory, which was directly under mine, to ask for names of investigators who

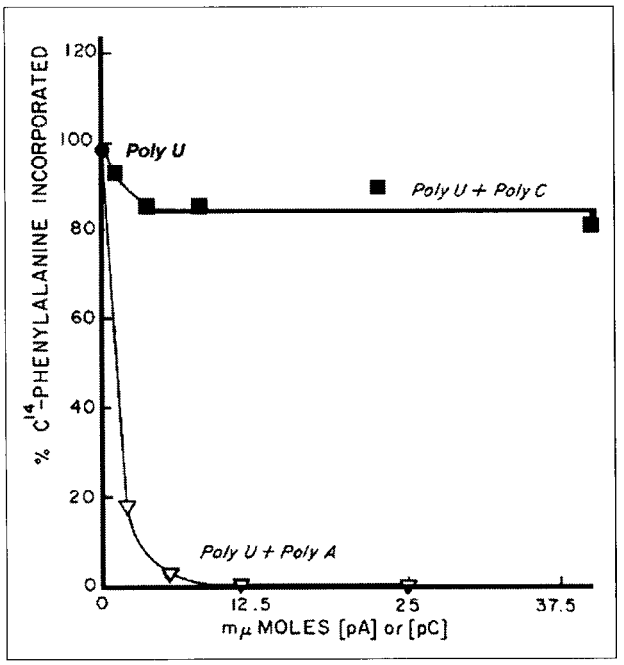

Figure 4. Addition of poly(A) completely inhibits the mRNA activity of poly(U) by the formation of double-stranded and triple-stranded helices. By contrast, addition of poly(C) has little effect on the mRNA activity of poly(U). This experiment, done in 1961, was the first the anti-sense RNA experiment [8].

Table 1. Characterization of ^{14}C-labeled polyphenylalanine, the synthesis of which is dependant upon poly(U): comparison of characteristics of product of reaction and poly-L-phenylalanine [8]

Treatment	Product of reaction	Poly-L-phenylalanine
6 N HCl for 8 h at 100°	Partially hydrolyzed	Partially hydrolyzed
12 N HCl for 48 h at 120–130°	Completely hydrolyzed	Completely hydrolyzed
Extraction with 33% HBr in glacial acetic acid	Soluble	Soluble
Extraction with the following solvents: H_2O, benzene, nitrobenzene, chloroform, *N,N*-dimethylformamide, ethanol, petroleum ether, concentrated phosphoric acid, glacial acetic acid, dioxane, phenol, ecetone, ethyl acetate, pyridine, acetophenone, formic acid	Insoluble	Insoluble

might have characterized polyphenylalanine. Michael Sela was the only person in the laboratory at the time; I knew that he worked with synthetic polypeptides so I asked if he knew anything about the solubility of polyphenylalanine. He said 'I do not know much, but I can tell you two things: one, polyphenylalanine is insoluble in most solvents; and second, it does dissolve in 15% hydrobromic acid dissolved in concentrated acetic acid.' I looked at him in delight as well as astonishment because I had never heard of such a solvent. Fifteen years later I learned that Michael Sela was the only person in the world who knew that polyphenylalanine dissolved in this esoteric solution because it is used to characterize C termini of proteins and he had mistakenly added it to polyphenylalanine, which, to his surprise, dissolved.

I was scheduled to give a talk in 1961 at the Vth International Congress of Biochemistry in Moscow. Just before leaving for Russia, I married Perola Zaltzman, a biochemist from Rio de Janeiro who worked with Sidney Udenfriend at the NIH, and we planned to meet for a leisurely, two week vacation after the meeting. I gave my talk in Moscow to an audience of ~35 people [11]. However, Francis Crick invited me talk again in a large symposium that he was chairing on nucleic acids, which I did to an extraordinarily enthusiastic audience. After returning to Bethesda, Fritz Lipmann generously gave me a partially purified transfer enzyme and we showed that phenylalanine-tRNA is an intermediate in the synthesis of polyphenylalanine directed by poly(U) [12]. A picture of Heinrich Matthei and me taken in 1962 is shown in Figure 5.

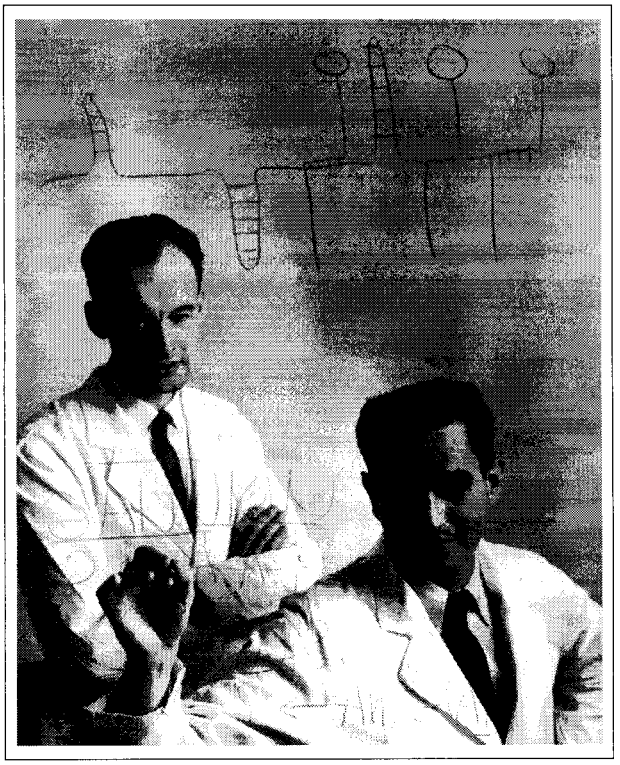

Figure 5. A picture of Heinrich Matthaei (left) and myself in 1962.

Soon afterwards I gave a talk at the Massachusetts Institute of Technology. A few years earlier Severo Ochoa from New York University had been awarded the Nobel Prize for his discovery, with Marianne Grunberg-Manago, of polynucleotide phosphorylase, which catalyzes the synthesis of randomly ordered polynucleotides. While I was answering questions from the audience, Peter Lengyel came up to the podium and told the audience that he and others in Ochoa's laboratory had used randomly ordered synthetic polynucleotides that contained several different nucleotide residues to direct the incorporation of other amino acids into protein. I flew back to Washington feeling very depressed because, although I had taken only two weeks to show that aminoacyl-tRNA is an intermediate in protein synthesis, I should have spent the time focusing on the more important problem of deciphering the genetic code. Clearly, I had to either compete with the Ochoa laboratory or stop working on the problem.

The next morning (Saturday) I went to the library in Leon Heppel's corridor to look up methods of synthesizing, purifying and characterizing randomly ordered polynucleotides whose synthesis was catalyzed by polynucleotide phosphorylase. Robert Martin (whose wife writes the *Miss Manners* newspaper column) was in the library and, when I told him what had happened, he suggested we synthesize randomly ordered polynucleotides that weekend. And that is exactly what we did. Bob, who is a superb, energetic investigator, stopped his own work and during the next few months synthesized and characterized many randomly ordered polynucleotides. Bob Martin played a major role in deciphering the genetic code.

Between 1961 and 1964 Bill Jones, Bob Martin and I determined the base compositions of RNA codons by directing amino acid incorporation into protein using many randomly ordered polynucleotide preparations that contained different combinations and proportions of bases [13–19]. In Figure 6 are the minimum species of bases required for mRNA codons. Although Severo Ochoa was a

Polynucleotides	Amino acids			
U	PHE			
C	PRO			
A	LYS			
G	—			
UC	LEU	SER		
UA	LEU	TYR	ILE	ASN
UG	LEU	VAL	CYS	TRP
CA	HIS	THR	GLN	ASN
CG	ARG	ALA		
AG	ARG	GLU		
UAG	ASP	MET		
CAG	ASP	SER		

Figure 6. The specificity of randomly ordered polynucleotide templates in stimulating amino acid incorporation into protein in *Escherichia coli* extracts. Only the minimum kinds of bases necessary for template activity are shown, so many amino acids that respond to randomly ordered polynucleotides composed of two or more kinds of bases are omitted. The base compositions of RNA codons were derived from these experiments [49]. Reproduced from Ref. [49].

fierce competitor during this time [20–28], we had never met. So, in 1961 I called him when I was in New York to arrange a meeting. I thought it would be more civilized to cooperate, or perhaps split the problem in some way, rather than compete with one another. Ochoa was very gracious, he invited me to his laboratory and introduced me his post-doctoral fellows, and we had tea in the library. However, there was no way we could collaborate. Later, to my horror, I found that I enjoyed competing. I focused on solving the problems that we were investigating and on working more effectively, rather than on winning or losing. The competition stimulated me to become more focused and I accomplished far more than I would have in its absence. From the beginning I vowed never to cut corners or reduce the rigor with which experiments were done to win the competition. Therefore, the quality of our work remained high throughout the deciphering process.

Years after the genetic code was deciphered, Mirko Beljansky told me that during a year's sabbatical in Ochoa's laboratory, Ochoa suggested that he see whether synthetic polynucleotides could direct cell-free protein synthesis in *E. coli* extracts. Beljansky spent a year trying to direct cell-free protein synthesis with poly(A), but was unsuccessful because the polylysine product of the reaction is a basic protein and is not precipitated by trichloracetic acid. Similarly, in Jim Watson's laboratory, Tissieres also tried to direct cell-free protein synthesis with poly(A), but did not detect the polylysine product for the same reason. Therefore, the idea of using synthetic polynucleotides to direct cell-free protein synthesis originated independently in at least three laboratories.

We showed that the code is formed of triplets by the amounts of radioactive histidine, threonine, asparagine, glutamine, lysine and proline that were incorporated into protein by five poly(A-C) preparations that contained different ratios of A and C [18,19] (Figure 7). The data for lysine and proline are not shown. The observed codon frequencies were compared with the theoretical frequencies of triplets or doublets in each poly(A-C) preparation calculated from the base composition determined for each polynucleotide preparation. The data showed that the codon for histidine is a triplet that contains one A and two Cs and the codons for asparagine and glutamine are triplets that contain two As and one C. The data also showed that two triplets or a doublet correspond to threonine.

A group picture of the people in my laboratory and their spouses taken in early 1964 is shown in Figure 8. There were ~20 people in my laboratory in all who deciphered the genetic code and about half of them (and their spouses) are in this picture. The occasion was a party for Brian Clark (middle of the first row in a suit), who was returning to Cambridge, England after several years in my laboratory. The post-doctoral fellows who came to my laboratory were superb and our work on deciphering the genetic code was very much a group project.

Although we and Ochoa's group had deciphered the nucleotide compositions of RNA codons, the nucleotide sequences were unknown. We tried several ways to solve this problem, but were unsuccessful. However, Kaji and Kaji [29] had shown that poly(U) stimulates the binding of radioactive polyphenylalanine-tRNA to ribosomes and Regina Cukier in my laboratory was using randomly ordered polynucleotides to direct the binding of aminoacyl-tRNA to ribosomes. I wondered how small a message one could use that would remain functional in directing

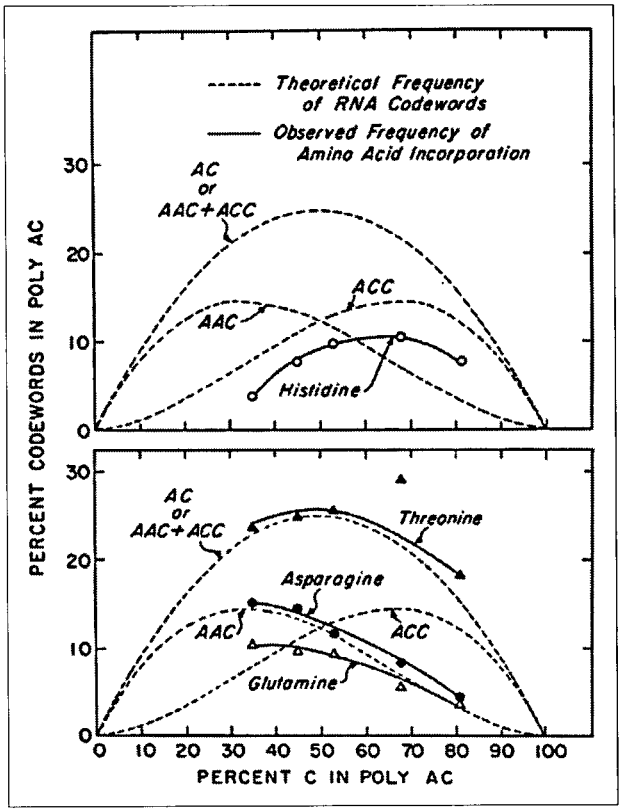

Figure 7. The theoretical frequencies of RNA codons in randomly ordered poly(AC) preparations that contain different proportions of A and C, compared with the observed frequencies of incorporation of radioactively labeled amino acids into protein. The codon for histidine contains one A and two Cs, and the codons for asparagine and glutamine contain two As and one C. These results showed that the code is a triplet code [18,19]. Reproduced from Ref. [18].

Figure 8. Going away party for Brian Clark who was returning to Cambridge, England after several years in the laboratory. The photograph was taken early in 1964. This is the only group photograph I have of the people in my laboratory while we were deciphering the genetic code. From left to right: Sid and Joan Pestka, Mrs. Marshall and Dick Marshall, Tom Caskey (partially hidden), Ty Jaouni, Mrs. Rottman, Norma Heaton (the marvelous technician with whom I worked for 39 years), Fritz Rottman, Brian Clark, Phil Leder, Shirley Shapiro (the secretary for the laboratory), Joel Trupin, Mrs. Trupin, myself, Bill Groves, Perola Nirenberg, Mrs. O'Neill and Charlie O'Neill.

aminoacyl-tRNA binding to ribosomes. Leon Heppel gave me a doublet, a triplet and a hexanucleotide that consisted only of U residues. I thought that ternary complexes, such as a ^{3}H-phenylalanine-tRNA–UUU trinucleotide–ribosome complex, would be retained by Millipore filters, whereas ^{3}H-phenylalanine-tRNA would be washed through the filter. The first experiment worked beautifully; the trinucleotide, UUU, stimulated binding of ^{3}H-phenylalanine-tRNA to ribosomes, but the doublet, UU, was without effect [30]. A similar experiment with oligo(A) preparations [49] is shown in Figure 9. It was clear that we could use this technique to determine the nucleotide sequences of RNA codons. However, most of the 64 trinucleotides had not been synthesized previously or isolated. Our major problem was to devise methods to synthesize trinucleotides of known sequence. Phil Leder saw an advertisement offering 0.5 g of each of the 16 doublets for $1500 per doublet. Although expensive (a few years earlier my salary for 6 months was $1500), we brought the lot. However, we received only 15 doublets because the US Customs Service confiscated one and used it all to test for drugs.

ADDITION	[^{14}C]-Lys-tRNA BOUND TO RIBOSOMES
	μmoles
ApA	0.01
ApApA	1.92
ApApApA	1.92
ApApApApA	1.92
ApApApApApA	2.71

Figure 9. Nucleotide sequences of RNA codons were determined by stimulating the binding to ribosomes of an appropriate species of radioactive aminoacyl-tRNA which recognized the trinucleotide. The aminoacyl-tRNA–trinucleotide–ribosome ternary complex was adsorbed on a Millipore filter and unbound aminoacyl-tRNA was removed by washing [30,49].

Marianne Grunberg-Manago was visiting Maxine Singer at that time, and both were experts on the use of polynucleotide phosphorylase. Phil Leder and Richard Brimacombe from my laboratory joined them in optimizing conditions for the polynucleotide phosphorylase catalyzed synthesis of oligoribonucleotides from doublet primers [31]. Trinucleotides then were separated from oligonucleotides of different chain lengths by electrophoresis.

In addition, in 1964, Leon Heppel (who was one of the best nucleic acid biochemists in the world at the time) suggested an esoteric method for the synthesis of trinucleotides and higher homologues, reported in a single sentence in one of his papers [32]. He showed that, in the presence of a high concentration of methanol, pancreatic RNase A catalyzes the synthesis of trinucleotides and higher homologues from oligoribonucleotide primers and pyrimidine 2′-3′-cyclic phosphates. We used both methods to synthesize triplets, as shown in Figure 10. Polynucleotide phosphorylase catalyzes the addition of nucleotide residues to the 3′-ends of doublets and other oligonucleotide primers, whereas pancreatic RNase A catalyzes the addition of nucleotides to the 5′ termini of oligonucleotide primers. Maxine Singer, Marianne Grunberg-Manago, and Leon Heppel, as well as Phil Leder, Mert Bernfield and Richard Brimacombe made major contributions to deciphering the genetic code. Bob Martin, Leon Heppel, and Maxine Singer are shown in Figure 11, and pictures of Phil Leder and Tom Caskey are shown in Figure 12.

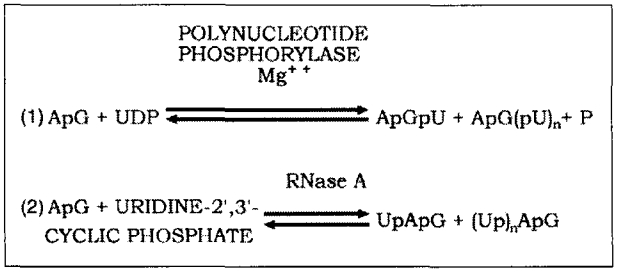

Figure 10. In 1964, most of the 64 possible trinucleotides had been neither synthesized nor isolated. The major problem was to devise methods for the synthesis of trinucleotides. The two enzymatic methods that we used are shown. (1) Maxine Singer (an expert on polynucleotide phosphorylase), Marianne Grunberg-Manago (who had discovered polynucleotide phosphorylase while in Severo Ochoa's laboratory), Philip Leder and Richard Brimacombe discovered conditions for the addition of one or a few nucleotide residues to the 3'-end of doublet primers, catalyzed by polynucleotide phosphorylase [31]. (2) Leon Heppel suggested an esoteric method for triplet synthesis that he, Whitfield and Markham had discovered in which pyrimidine-2', 3'-cyclic phosphates are added to the 5'-ends of doublet primers catalyzed by pancreatic RNase A in the presence of a high concentration of methanol [32]. Merton Bernfield synthesized half of the trinucleotides in our laboratory using this method.

Figure 12. Phil Leder came to my laboratory as a post-doctoral fellow and played a major role in deciphering the genetic code. He was the first to decipher the nucleotide sequence of a codon [33]. Tom Caskey was another post-doctoral fellow in my laboratory. Together with Dick Marshall, he compared the genetic code of *Escherichia coli* with that of *Xenopus* and hamsters and showed that the code is universal [41]. Later, Tom Caskey and his colleagues worked on the mechanism of termination of protein synthesis.

It took my colleagues, Philip Leder, Merton Bernfield, Joel Trupin, Sid Pestka, Fritz Rottman, Richard Brimacombe, Charles O'Neal, French Anderson, Don Kellogg, Ty Jaouni and myself a year to synthesize the 64 trinucleotides and test each against 20 different radioactive aminoacyl-tRNA preparations to decipher the nucleotide sequences of RNA codons [30,33,35–40]. Philip Leder deciphered the first nucleotide sequence of a valine RNA codon using trinucleotide-dependent binding of aminoacyl-tRNA to ribosomes [33]. We soon found that the third bases of synonym RNA codons varied systematically [33,35–40] and, using purified aminoacyl-tRNA fractions [39–41], we found four patterns of degeneracy, as shown in Figure 13. Francis Crick coined the term 'wobble' to describe the third-base degeneracy of RNA codons [42]. Generously, Robert Holley gave us a highly purified fraction of yeast alanine-tRNA that he and his colleagues had sequenced in what was the first determination of the nucleotide sequence of a nucleic acid [43]. Philip Leder and I showed that this species of alanine-tRNA recognizes the codons GCU, GCC and GCA, and that the nucleotide residue in the tRNA anti-codon that recognizes U, C or A in the third position is hypoxanthine [40].

After we had published the nucleotide sequences of 54 of the 64 RNA codons [30,33,35–40] the great nucleic acid chemist, Gobind Khorana and his colleagues reported the nucleotide sequences of RNA codons determined using chemically synthesized trinucleotides [44]. The genetic code is shown in Figure 14. Clark and Marcker [45] showed that *N*-formyl-methionine-tRNA initiates protein synthesis by recognizing the codon AUG, whereas AUG at internal positions corresponds to methionine residues. UAA, UAG and UGA were reported to correspond to the termination of protein synthesis by Sidney Brenner and Alan Garen and their colleagues [47–48]. The hydrophobic amino acids Phe, Leu, Ile, Met and Val correspond to chemically similar codons that have U as the second base. By contrast, the hydrophilic amino acids Tyr, His, Gln, Asn, Lys, Asp and Glu correspond to codons with A as the second base. In addition, amino acids with chemically similar side chains, such as Asp and Glu, and Asn and Gln, have chemically similar codons. Clearly, the arrangement of codons and amino acids is not random.

After deciphering the genetic code we asked the question 'is the code universal?'. Thomas Caskey and

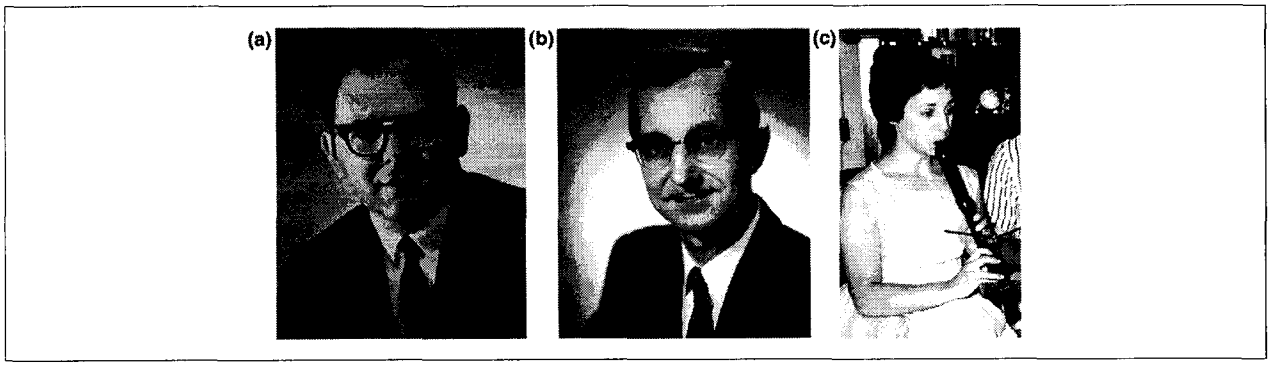

Figure 11. Three investigators at the NIH played major roles in deciphering the genetic code. Bob Martin synthesized and characterized many randomly ordered polynucleotides had a very important role in deciphering the base compositions of RNA codons. Leon Heppel was one of the few nucleic acid biochemists in the world at that time. He gave me compounds and advice when I needed it, and suggested the use of pancreatic RNase A to catalyze trinucleotide and higher homologue synthesis, a method that he had discovered earlier [32]. Maxine Singer came to the NIH as a post-doctoral fellow working with Leon Heppel. She was an expert on polynucleotide phosphorylase and helped devise conditions for the synthesis of trinucleotides catalyzed by polynucleotide phosphorylase [31].

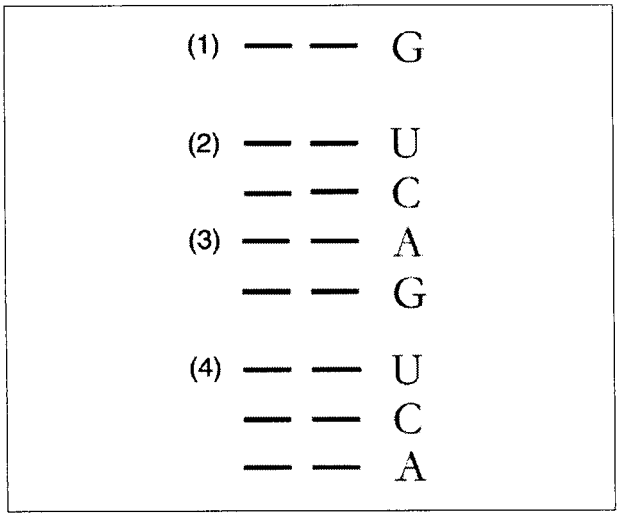

Figure 13. Four kinds of degeneracy for synonym codons were found using purified aminoacyl-tRNA preparations. Robert Holley gave us a highly purified preparation of yeast Ala-tRNA, which was the first nucleic acid to be sequenced [43], and we showed that it recognized the codons GCU, GCC and GCA [40]. Hypoxanthine in the t-RNA anticodon recognizes U, C or A in the third position of RNA codons.

Richard Marshall purified tRNA fractions from *Xenopus* embryos and hamster liver. Comparing the genetic code of *E. coli* to that of *Xenopus* and hamster, we found that the code is essentially universal [41]. These results had a profound philosophical impact on me because they indicate that all forms of life on this planet use essentially the same language. Some dialects have been reported subsequently in some organisms, but all are modifications of the same genetic code.

We deciphered the genetic code over a period of about five years, from 1961 to 1966. This was a group project and the post-doctoral fellows in my laboratory during this period (Table 2) contributed in many important ways. In addition, Robert Martin, Leon Heppel, Maxine Singer and Marianne Grunberg-Manago played major roles in deciphering the genetic code, and Ochoa and Khorana and their colleagues also contributed to the deciphering of the genetic code. Deciphering the genetic code was the first project that I worked on as an independent investigator, and it was an extraordinarily exciting, fun-filled project to explore and solve. Although many excellent problems related to the code and protein synthesis remained after the code was deciphered, I decided to switch to the more challenging field of neurobiology, a field I am still exploring.

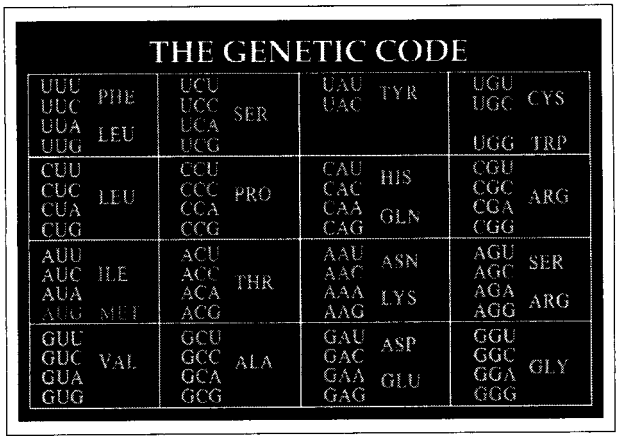

Figure 14. It took us about a year to synthesize the 64 trinucleotides and test each against 20 radioactive aminoacyl-tRNA preparations to determine the nucleotide sequences of RNA codons [30,33,35–40]. Gobind Khorana and his colleagues synthesized the 64 trinucleotides chemically and also determined nucleotide sequences of some RNA codons [44]. The green AUG corresponds to methionine and *N*-formyl-methionine tRNA, an initiator of protein synthesis [45]. The red codons specify the termination of protein synthesis [46–48].

Table 2. Post-doctoral fellows in my laboratory 1961–1966. Deciphering the genetic code was very much a group affair and many of these individuals contributed greatly to this work. There were nine or fewer fellows at any one time, but not all worked on the code

Heinrich Matthaei	Thomas Caskey
Oliver W. Jones	Joseph Goldstein
Samuel Barondes	Edward Scolnick
Philip Leder	Richard Tomkins
Sidney Pestka	Arthur Beaudet
Fritz Rottman	French Anderson
Merton Bernfield	Charles O'Neal
Brian Clark	William Groves
William Sly	Raymond Byrne
Judith Levin	Regina Cukier
Joel Trupin	Donald Kellogg
Richard Brimacombe	Michael Wilcox
Richard Marshall	Frank Potugal
Dolph Hatfield	

Acknowledgements

I thank Michael Spencer for his help in digitizing and formatting the figures and tables.

References

1 Nirenberg, M.W. and Jakoby, W.B. (1960) Enzymatic utilization of γ-hydroxybutyric acid. *J. Biol. Chem.* 235, 954–960
2 Pollock, M.R. (1962) Exoenzymes. In *Bacteria* (4) (Grinsalus, I.G. and Stanier, R.Y., eds), pp. 121–178
3 Hershey, A.D. *et al.* (1953) Nucleic acid economy in bacteria infected with bacteriophage T2. I. Purine and pyrimidine composition. *J. Gen. Physiol.* 36, 777–789
4 Volkin, E. and Astrachan, L. (1956) Phosphorus incoporation in *Escherichia coli* ribonucleic acid after infection with bacteriophage T2. *Virology* 2, 149–161
5 Lamborg, M.R. and Zamecnik, P.C. (1960) Amino acid incorporation into protein by extracts of *E. coli*. *Biochim. Biophys. Acta* 42, 206–211
6 Matthaei, J.H. and Nirenberg, M.W. (1961) The dependence of cell-free protein synthesis in *E. coli* upon RNA prepared from ribosomes. *Biochem. Biophys. Res. Commun.* 4, 404–408
7 Matthaei, J.H. and Nirenberg, M.W. (1961) Characteristics and stabilization of DNase sensitive protein synthesis in *E. coli* extracts. *Proc. Natl. Acad. Sci. U. S. A.* 47, 1580–1588
8 Nirenberg, M.W. and Matthaei, J.H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* 47, 1588–1602
9 Kameyama, T. and Novelli, G.D. (1960) The cell-free synthesis of β-galactosidase by *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 2, 393–396
10 Tissieres, A. *et al.* (1960) Amino acid incorporation into proteins by *Escherichia coli* ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* 46, 1450–1463
11 Nirenberg, M.W. and Matthaei, J.H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic template RNA. In *Biological Structure and Function at the Molecular Level* (Engelhardt, V.A., ed.), pp. 184–189, The MacMillan Co
12 Nirenberg, M.W. *et al.* (1962) An intermediate in the biosynthesis of

polyphenylalanine directed by synthetic template RNA. *Proc. Natl. Acad. Sci. U. S. A.* 48, 104–109
13 Martin, R.G. *et al.* (1961) Ribonucleotide composition of the genetic code. *Biochem. Biophys. Res. Commun.* 6, 410–414
14 Matthaei, J.H. *et al.* (1962) Characteristics and composition of RNA coding units. *Proc. Natl. Acad. Sci. U. S. A.* 48, 666–677
15 Jones, O.W. and Nirenberg, M.W. (1966) Degeneracy in the amino acid code. *Biochim. Biophys. Acta* 119, 400–406
16 Nirenberg, M.W. *et al.* (1963) Approximation of genetic code via cell-free protein synthesis directed by template RNA. *Fed. Proc.* 22, 55–61
17 Nirenberg, M.W. and Jones, O.W. Jr (1963) The current status of the RNA code. In *Symposium on Informational Macromolecules* (Vogel, H., ed.), pp. 451–465, Academic Press
18 Nirenberg, M.W. *et al.* (1963) On the coding of genetic information. *Cold Spring Harb. Symp. Quant. Biol.* 28, 549–557
19 Jones, O.W. and Nirenberg, M.W. (1966) Degeneracy in the amino acid code. *Biochim. Biophys. Acta* 119, 400–406
20 Lengyel, P. *et al.* (1961) Synthetic polynucleotides and the amino acid code. *Proc. Natl. Acad. Sci. U. S. A.* 47, 1936–1942
21 Speyer, J.F. *et al.* (1962) Synthetic polynucleotides and the amino acid code, II. *Proc. Natl. Acad. Sci. U. S. A.* 48, 63–68
22 Lengyel, P. *et al.* (1962) Synthetic polynucleotides and the amino acid code, III. *Proc. Natl. Acad. Sci. U. S. A.* 48, 282–284
23 Speyer, J.F. *et al.* (1962) Synthetic polynucleotides and the amino acid code, IV. *Proc. Natl. Acad. Sci. U. S. A.* 48, 441–448
24 Basilio, C. *et al.* (1962) Synthetic polynucleotides and the amino acid code, V. *Proc. Natl. Acad. Sci. U. S. A.* 48, 613–616
25 Wahba, A.J. *et al.* (1962) Synthetic polynucleotides and the amino acid code, VI. *Proc. Natl. Acad. Sci. U. S. A.* 48, 1683–1686
26 Gardner, R.S. *et al.* (1962) Synthetic polynucleotides and the amino acid code, VII. *Proc. Natl. Acad. Sci. U. S. A.* 48, 2087–2094
27 Wahba, A.J. *et al.* (1963) Synthetic polynucleotides and the amino acid code, VIII. *Proc. Natl. Acad. Sci. U. S. A.* 49, 116–122
28 Speyer, J.F. *et al.* (1963) Synthetic polynucleotides and the amino acid code. *Cold Spring Harb. Symp. Quant. Biol.* 28, 559–567
29 Kaji, A. and Kaji, H. (1963) Specific interaction of soluble RNA with polyribonucleic acid induced polysomes. *Biochem. Biophys. Res. Commun.* 13, 186–192
30 Nirenberg, M. and Leder, P. (1964) RNA codewords and protein synthesis. I. The effect of trinucleotides upon the binding of sRNA to ribosomes. *Science* 145, 1399–1407
31 Leder, P. *et al.* (1965) Synthesis of trinucleoside diphosphate with polynucleotide phosphorylase. *Biochemistry* 4, 1561–1567
32 Heppel, L.A. *et al.* (1955) Nucleotide exchange reactions catalyzed by ribonuclease and spleen phosphodiesterase. 2. Synthesis of polynucleotides. *Biochem. J.* 60, 8–15
33 Leder, P. and Nirenberg, M.W. (1964) RNA codewords and protein synthesis. II. Nucleotide sequence of a valine RNA codeword. *Proc. Natl. Acad. Sci. U. S. A.* 52, 420–427
34 Nirenberg, M.W. (1964) Protein synthesis and the RNA code. *Harvey Lectures Series 59*, pp. 155–185
35 Bernfield, M.R. and Nirenberg, M.W. (1965) RNA codewords and protein synthesis. IV. The nucleotide sequences of multiple codewords for phenylalanine, serine, leucine and proline. *Science* 147, 479–484
36 Trupin, J.S. *et al.* (1965) RNA codewords and protein synthesis. VI. On the nucleotide sequences of degenerate codeword sets for isoleucine, tyrosine, asparagine, and lysine. *Proc. Natl. Acad. Sci. U. S. A.* 53, 807–811
37 Nirenberg, M.W. *et al.* (1965) RNA codewords and protein synthesis. VII. On the general nature of the RNA code. *Proc. Natl. Acad. Sci. U. S. A.* 53, 1161–1168
38 Brimacombe, R. *et al.* (1965) RNA codewords and protein synthesis. VIII. Nucleotide sequences of synonym codons for arginine, valine, cysteine, and alanine. *Proc. Natl. Acad. Sci. U. S. A.* 54, 954–960
39 Kellogg, D.A. *et al.* (1966) RNA codons and protein synthesis. IX. Synonym codon recognition by multiple species of valine-, alanine-, and methionine-sRNA. *Proc. Natl. Acad. Sci. U. S. A.* 55, 912–919
40 Nirenberg, M.W. *et al.* (1966) The RNA code and protein synthesis. *Cold Spring Harb. Symp. Quant. Biol.* 31, 11–24
41 Marshall, R.E. *et al.* (1967) RNA codewords and protein synthesis. XII. Fine structure of RNA codewords recognized by bacterial, amphibian, and mammalian transfer RNA. *Science* 155, 820–826
42 Crick, F.H.C. (1966) Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* 19, 548–555
43 Holley, R.W. *et al.* (1965) Structure of a ribonucleic acid. *Science* 147, 1462–1465
44 Söll, D. *et al.* (1965) Studies on polynucleotides, XLIX. Stimulation of the binding of aminoacyl-sRNA's to ribosomes by ribotrinucleotides and a survey of codon assignments for 20 amino acids. *Proc. Natl. Acad. Sci. U. S. A.* 54, 1378–1385
45 Clark, B.F.C. and Marcker, K.A. (1966) The role of N-formyl-methionyl-sRNA in protein biosynthesis. *J. Mol. Biol.* 17, 394–406
46 Brenner, S. *et al.* (1965) Genetic code: the 'nonsense' triplets for chain termination and their suppression. *Nature* 206, 994–998
47 Weigert, M.G. and Garen, A. (1965) Base composition of nonsense codons in *E. coli* evidence from amino-acid substitutions at a tryptophan site in alkaline phosphatase. *Nature* 206, 992–994
48 Brenner, S. *et al.* (1967) UGA: A third nonsense triplet in the genetic code. *Nature* 213, 449–450
49 Nirenberg, M. (1969) The genetic code. In *Les Prix W. Nobel en 1968. Nobel Foundation* (Stockholm, P.A. *et al.*, eds), pp. 221–241, Norstedt and Sőner

News & Features on *BioMedNet*

Start your day with *BioMedNet's* own daily science news, features, research update articles and special reports. Every two weeks, enjoy *BioMedNet Magazine*, which contains free articles from *Trends, Current Opinion, Cell* and *Current Biology*. Plus, subscribe to Conference Reporter to get daily reports direct from major life science meetings.

http://news.bmn.com